



General Data Analysis and Visualization using HPC

The University of Tennessee Center for Remote Data Analysis & Visualization (RDAV)

Jian Huang
Associate Professor | EECS
Associate Director | RDAV
University of Tennessee



Funded by NSF Teragrid XD, RDAV is a joint effort by University of Tennessee, National Center for Supercomputing Applications (NCSA), Oak Ridge National Laboratory, Lawrence Berkeley National Laboratory and University of Wisconsin.



Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation

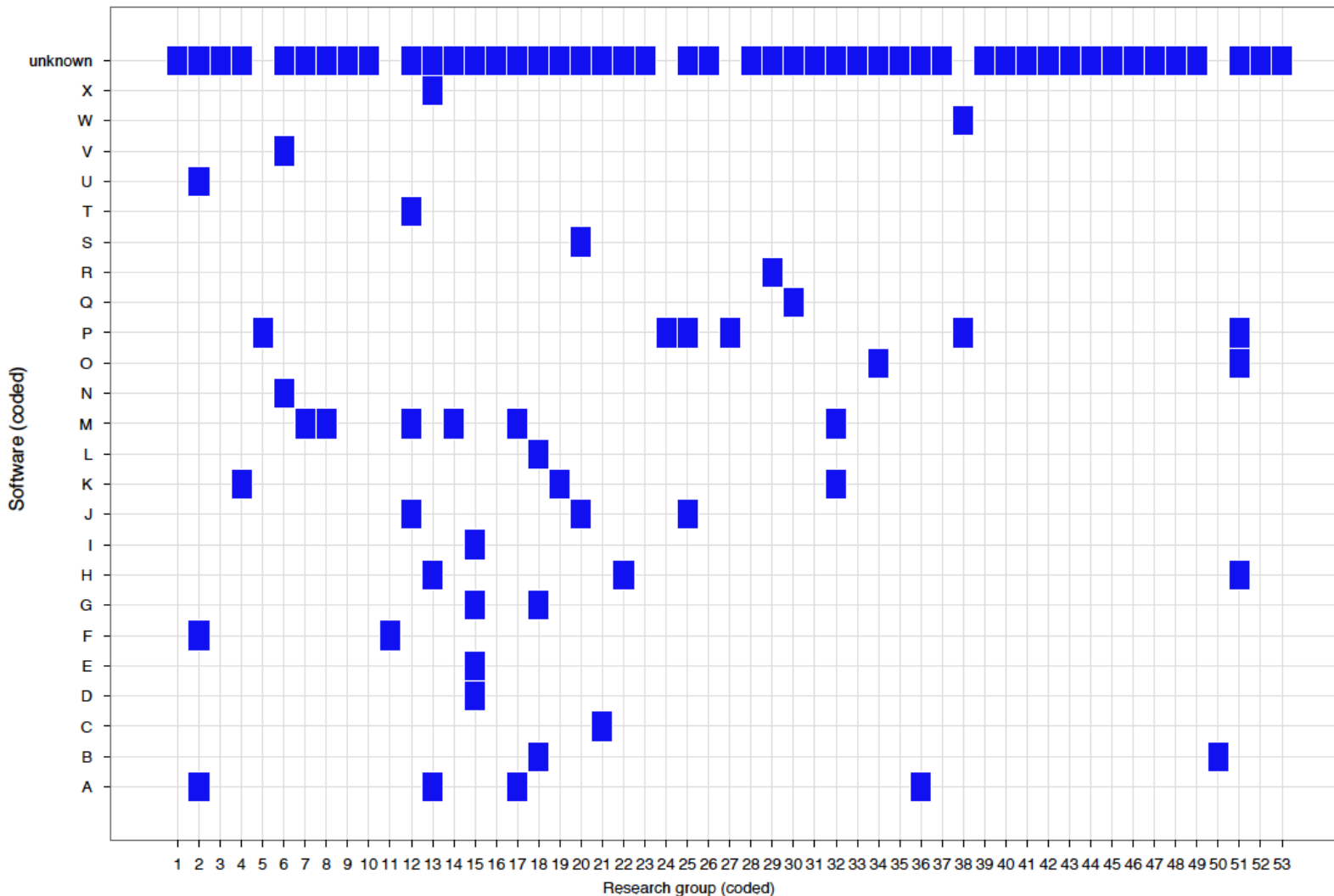


RDAV - General Data Analysis and Visualization Services

- Located at National Institute of Computational Science
- Goal
 - To provide data analysis and visualization services that enable computation and data intensive science
- Approach
 - Unique HPC system – SMP: 1024 core, 4TB memory
 - Comprehensive expertise



Diverse Domain Specific Software



Granularity of Parallelism

Parameter sweep using serial apps:
MaxEnt, Energy+
...

Parallel apps:
Paraview, VisIt

Parallel programming: R, Python, Java, Matlab

MPI, pthreads



Scalable News Mining

Scaling up news mining code in Perl

Used RDAV/UT's in-house developed glue software – Eden (released as rdav_eden on sourceforge.net)

Quoting from the NCSA user Kalev Leetaru's emails:

- .. Scott, this is a ***fantastic*** tool, exactly what I've been looking for!....
- .. I will say that Eden is a great great tool, it is proving ***tremendously*** useful and really helping to speed things up for me...

The image shows two overlapping web browser windows. The top window is the BBC News Technology page, dated 9 September 2011, with the headline 'Supercomputer predicts revolution'. The article discusses how feeding a supercomputer with news stories could help predict major world events, based on a study charted by d... The bottom window is the Nature News page, dated 13 September 2011, with the headline 'News mining might have predicted Arab Spring'. This article, by Phillip Ball, discusses how a new study shows how data mining of news reportage can reveal the possibility of future crises well before they happen. It mentions computer scientist Kalev Leetaru at the University of Illinois at Urbana-Champaign, who has trawled through a vast collection of news reporting and examined the 'tone' of the news about Tunisia, Egypt, and Libya. The Nature News page also includes sections for 'Stories by subject' (History, Policy, Mathematics), 'Stories by keywords' (Culturomics, Data mining, Conflict, War, Social sciences, News), and 'This article elsewhere' with links to various social media and bookmarking services. A collage of newspaper front pages related to the Arab Spring is visible in the bottom right corner of the Nature News window.

Visualizing the Magnetosphere

Electron and ion vorticity from the
asymmetric VPIC simulation

120GB per vector field per timestep.

Visualized using a discrete particle
noise reduction technique deployed
in ParaView.

**Collaboration with Homa
Karimabadi (UCSD/SciberQuest).
Visualization by RDAV/LBNL.**

